REVIEW OF COLUMBIA RIVER TEMPERATURE ASSESSMENT: SIMULATION METHODS

JUNE 1999

Prepared for:

Potlatch Corporation
Lewiston, Idaho

Prepared by:

Peter Shanahan, Ph.D, P.E.
M. Bruce Beck, Ph.D.

HAI HydroAnalysis, Inc.

481 Great Road, Suite 3
Acton, Massachusetts 01720
(978) 263-1092
fax: (978) 263-8910
e-mail: shanahan@ma.ultranet.com

REVIEW OF
COLUMBIA RIVER TEMPERATURE ASSESSMENT: SIMULATION METHODS

INTRODUCTION

This document provides a review of the report "Columbia River Temperature Assessment: Simulation Methods" prepared by John Yearsley of U.S. Environmental Protection Agency Region 10 in February 1999. The report employs sophisticated modeling and parameter estimation techniques to assess temperature in the Columbia River system and place bounds on the uncertainty in predicting temperature.

The introduction to the EPA report lists the following three sources that may contribute to changes in the temperature of the Columbia and Snake Rivers: impoundments, hydrologic modifications, and watershed modifications. The report does not consider a potentially significant contributor to temperature change, global climate change. The report indicates that it is the objective of the study "to assess the relative importance of these sources with respect to changes in the temperature regime of the main stem Columbia River in Washington and Oregon and in the Snake River in Washington." This objective is partially accomplished. The report limits its focus to the effects of dams and a blanket consideration of contributing tributary temperature.

The number of simulations in the study and their sophistication are insufficient to achieve fully the stated objectives and the study should be viewed as a screening level analysis to identify potential factors that affect river temperature. The report notes on page 5 that the model is indeed a screening model, designed only "to identify critical areas for additional analysis." Thus, the model results should be recognized to be approximate and exploratory rather than definitive. This report would be an inappropriate basis for policy decisions other than the identification of areas for further research.

MATHEMATICAL MODEL DEVELOPMENT

The author elected to write a new computer program for modeling temperature rather than relying upon an established model such as QUAL-2E (Brown and Barnwell, 1987). As explained by author, his approach provides some legitimate advantages for the problem in question. Nonetheless, the failure to use an established model that has experienced substantial use and prior review places an added burden on the author to verify his model.

Typically, computer-model verification is accomplished by using the model to simulate one or more problems for which there is a well-known analytical solution. Close agreement between the model solution and the analytical solution "verifies" the model; that is, it demonstrates that the model accurately solves the type of problem that it was designed to solve (National Research Council, 1989, pp.235-236). The greater the number of such test cases performed, the more confident one can be of the model's accuracy and validity. In essence, verification is a process to guarantee against erroneous computer code. It protects against the situation in which erroneous code begets erroneous results, known in computer science by the shorthand, "garbage in, garbage out."

No verification of the Columbia River temperature model was provided in the reviewed document. Use of the model for screening purposes does not obviate the need for verification. Without such verification, the model's accuracy cannot be assessed and the model cannot be relied upon. The U.S. EPA should be requested to provide documentation of model verification in accordance with accepted standards.

While potential programming errors could compromise the model's validity, the model has a solid technical foundation. The model equations as presented in equations 1, 2, 4, 13-15, and 17-21 are well known and accepted. The equations for stream temperature are solved in the EPA model using a mixed Eulerian-Lagrangian solution scheme. This approach is much less common in surface-water-quality modeling than an Eulerian approach, but is nonetheless entirely valid. The Eulerian approach, which is used for example in QUAL2E, solves for temperature at fixed points along the river. In essence, the Eulerian solution looks at the entire river at a single time. In the Lagrangian approach, the numerical solution follows a particular parcel of water down the river. In essence, this technique looks at this single parcel, following it in time and space as it moves downriver with the streamflow. Travel time for the parcel, τ, and its location along the river, x, are directly related by the stream velocity, U: $x = U\tau$. The Lagrangian approach takes advantage of this relationship to eliminate the distance variable from the equations being solved.

The mixed solution in the Columbia River model uses a Lagrangian algorithm to model the effect of water flowing downstream in the river, but an Eulerian algorithm to model mixing. The main advantages of the mixed Lagrangian-Eulerian technique, in general, are accurate representation of dispersion and the ability to model sharp transitions in concentration. These are not factors in the current model version since diffusion is not modeled and the concentration (temperature) changes gradually. Thus, although these factors are cited on page 10 of the report, the real advantage of using this solution in the Columbia River model seems to be that it simplifies the accompanying Kalman filtering analysis of model parameters. The role of Kalman filtering in the Columbia River model is further described below.

DATA SOURCES AND PARAMETER ESTIMATION

A multi-step process that considered the statistical properties of temperature prediction and measurement and their uncertainty was used to establish the various parameters in the model. First, model parameters were set based directly on the available data and literature. This first step is called the deterministic parameter estimation. Next, these parameters were adjusted to produce results deemed "unbiased." In essence, this is the same as the manual calibration process used in traditional water-quality modeling. Finally, the variance of the systems model (i.e., the model for river temperature) was estimated as the third and final step of the parameter estimation process. Variance is a statistical measure of the variation of the model results around its mean prediction and depends upon the error in both the model and field data.

The Columbia River model requires a large number of input data to represent the river's geometry and the basin hydrology and meteorology. As discussed on page 13 of the EPA report, numerous past studies of this river system provide an unusually rich database for modeling. The first step in the model development, deterministic parameter estimation, was thus a fairly straightforward and familiar process. A few exceptions are discussed below.

On page 14, the report describes the basis for estimating the river's hydraulic parameters as a function of streamflow under the dam removal scenario. These estimates were made for flows of 60,000, 120,000 and 240,000 cfs in the Snake River. Unfortunately, the minimum flow of 60,000 cfs exceeds the average monthly flow in the Snake River below Ice Harbor Dam for the months of July through March and is more than twice the monthly average flows for August, September, and October. The model hydraulic parameters thus may be inappropriate for modeling the dam removal scenario during the critical low-flow months. This is significant because these months are also the most likely to experience temperature exceedances.

In computing surface heat transfer, the author has appropriately opted to use a heat budget model when simpler, but less accurate, approaches are sometimes used. The heat budget model, which is presented on page 15 and in equations 17 through 21, is based on an authoritative reference (Wunderlich and Gras, 1967) although a somewhat later and more accessible version of the same reference exists (TVA, 1972). The presentation in the EPA report is incomplete, however, in that the report fails to define all variables and units in the equations. This creates some uncertainty as to exactly what relations were actually used. One critical factor in calculating surface heat transfer involves the evaporative heat flux in equation 19. There are multiple formulas for the dependence of this flux on wind speed and the formula chosen may significantly alter the computed heat transfer (Shanahan, 1985). The EPA report fails to clarify the exact formulation used. With respect to the shortwave radiation in equation 17, the atmospheric attenuation coefficient has been shown in some studies in the Pacific Northwest to take on unusual values owing to haze (Findikakis *et al.*, 1980). This may be a factor in some areas of the watershed, which would necessitate special parameter selection.

On page 15, the report states that **daily** temperature values are not always available for the upstream stations that form the model boundaries. The report then presents, on page 16, an empirical equation for **weekly** temperature that was used to fill in missing temperature values. The mismatch between the daily and weekly periods is significant in that there may be a significant time lag between meteorological conditions and the resulting stream temperatures. As a result, the relation for weekly conditions is likely to be substantially different from the relation for daily conditions. It is unclear from the report whether daily temperature values were actually derived or whether weekly temperatures were used during periods of missing data. In either case, there is substantial margin for error in fixing the temperature at the model inflow points.

Under the subheading "Systems Model Bias and Error" starting on page 17, the report describes how Kalman filtering was used. Kalman filtering augments the supposedly certain prediction made by a traditional water-quality model with a probabilistic prediction that recognizes multiple sources of uncertainty. In a traditional model, the temperature at some time step k is determined from the temperature at the preceding time step k-1, using the model equations and known parameters. In fact, however, the temperature at time step k-1 is known imperfectly, the parameters for stepping from time step k-1 to time step k are uncertain, and even the model formulation itself probably has errors. If the model results are compared with field measurements, the field measurements must also be recognized as having some error. The Kalman filtering approach recognizes these various sources of error and incorporates them into the model formulation. The result is an estimate of the uncertainty in the model predictions that can be used to help guide the calibration procedure.

Kalman filtering is a complicated and specialized technique. Accordingly, expert review of the EPA's application of Kalman filtering in the Columbia River temperature analysis was sought from Professor M. Bruce Beck of the University of Georgia. Dr. Beck is an internationally-known specialist in the application of Kalman filtering to surface-water quality. Dr. Beck's review is appended to this review. Dr. Beck finds no fault in the technical aspects of the Kalman filter analysis, but raises some cautions as to the interpretation of the results. These cautions are pointed out in the discussion that follows.

The EPA study used Kalman filtering in an approach that closely follows that presented by Van Geer *et al.* (1991) for a ground-water modeling application. Despite the change in environmental medium, the approach remains valid for the EPA application. As presented by Van Geer *et al.*, Kalman filtering provides information on the uncertainty of the model prediction and helps guide the calibration process. According to Van Geer *et al.*, one can achieve a better calibration using Kalman filtering than the traditional, deterministic approach. In a personal communication, Dr. Beck has indicated his strong disagreement with this assertion.

The Kalman filtering procedure is complicated, as is implied in equations 5 through 12 on page 9, and the description in the EPA report is spare and difficult to follow. Accordingly, the procedure is summarized in the next three paragraphs, which are based on the more lucid description in the paper by Van Geer *et al.* (1991). These three paragraphs, which are quite technical, can be skipped without loosing the overall sense of this review. With respect to the procedure described in the EPA report, equations 8 and 12 include errors. In equation 8, the second instance of f_{k-1} should instead be its transpose, f_{k-1}^{T}. In equation 12, $\underline{v}_k$ and $\underline{z}_k$ are vectors and should both be underscored.

The Kalman filtering process, as described in equations 5 through 12, marches through time in discrete time steps. It consists of two sub-steps at each time step: first a prediction is made strictly from the model equations, and second it is corrected based on the measured data.

The first sub-step is the prediction. At each time step k in the temperature simulation, the temperature at the various measurement locations along the river (represented by the vector $\underline{T}_k$) is predicted with equation 7 as a function of the system matrix (f_{k-1}) and the temperature at the last time step, $\underline{T}_{k-1}$. The system matrix is simply the temperature equations 2 and 4 in another form. In parallel with the temperature prediction at time step *k*, the uncertainty in the prediction is estimated with equation 8. The uncertainty in the temperature is a matrix, P_k, in which the diagonal elements are the variances of each temperature value (i.e., the temperature at each station) and the off-diagonal elements are the covariances between the temperatures at different stations. This matrix is known as the error covariance. Like the temperature vector, the error covariance matrix is predicted based on its value at the last time step.

Following the predictor sub-step is the corrector or update sub-step. Here, the predicted temperature is updated with the actual temperature measurements, $\underline{z}_k$, using equation 9. Equation 9 is simply a weighted average of the predicted temperature and the measured temperature, but with the weighting changing as the simulation progresses. The weighting is captured in the so-called Kalman gain, K_k, which is also a matrix and is calculated in equation 11. In the actual computational sequence, equation 11 is completed before equations 9 and 10. The error covariance is similarly updated in equation 10. At the end of the update sub-step, the calculation for time step *k* is completed and the process begins again with the predictor sub-step for time step *k+1*.

An outcome from Kalman filtering is the innovations sequence, equation 12, which shows the error between the measured and predicted temperature at each location along the river at each time step *k* in the simulation. The goal in calibrating the deterministic temperature model is to adjust the model so as to minimize the mean of the values of this error term over time. In addition, the error covariance matrix (the Σ_Q term in equations 5 and 8) is adjusted until the innovations sequence satisfies certain statistical properties discussed below. Varying, one at a time, the deterministic model and the properties of the stochastic model error adjusts the model. As described in the EPA report, the only deterministic parameter varied was the meteorological data station assigned to each reach of the river. The assignment of stations was varied manually until, according to the report, "the mean of the innovations vector was small." No specific description of "small" is given, although Figures 6 through 13 allow a visual evaluation of the error. Dr. Beck cautions that Figures 6 through 13 appear to compare updated temperature predictions, and thus may present a more favorable comparison to the field data than is appropriate. As indicated in Dr. Beck's review, the exact character of the simulated values in Figures 6 through 13 should be clarified.

The stochastic error term that was varied is the estimate of the error in the system model. This error is represented by $\underline{w}_{k-1}$ in equation 5, and it is assumed to be a Gaussian distribution with zero mean and variance Σ_Q. This error is not known at the start of the modeling exercise, so it is given an initial guess and then corrected by trial-and-error based on the results of the Kalman filtering. The corrections entail changing the values of the assumed statistical variance matrix, Σ_Q. The stochastic part of the model is determined to be calibrated when the values of Σ_Q cause the model error computed from the innovations sequence to match the theoretical error predicted by the model. Mechanically, this match is computed using equations 23 and 24 on page 17. As Dr. Beck points out in his review, the values of Σ_Q are expected to differ between simulations of the existing situation with dams and predictions for a future situation without dams. However, it appears that the same values of Σ_Q were used for both scenarios. Dr. Beck also points out that assumptions made concerning the character of covariance terms in Σ_Q are inadequately discussed in the report.

The model results raise some questions. The text indicates that data were available for the period 1975 through 1995, but calibration results for only 1990 through 1995 are shown. It cannot be determined from the report whether the entire period of record was used to calibrate the model or if only the 1990-1995 subset of the record was used. It would not be inappropriate to base the calibration on the 1990-1995 period only, since page 16 indicates the data are more reliable then, but the data selection should be clarified.

The innovations sequence is a measure of the difference between the temperature predicted by the model and that actually measured. The innovations sequence is shown in Figures 14 through 21 over the calibration period at a number of measurement stations along the rivers. The error is relatively large—greater than 3 or 4 degrees. Moreover, the figures plot a 30-day moving average, implying that some daily values are even more in error. The report is deficient in explaining the meaning, significance, and limitations of these results. The figures illustrate the calibration of the deterministic model where the goal is to get the mean of all plotted values to equal zero. This can be equivalently thought of as getting the area under the plotted curve above the x-axis (0-degree line) equal to the corresponding area below the x-axis. It appears that at some of the stations, the calibration fell well short of this goal. The peaks and valleys in Figures 14 through 21 indicate that the model appears consistently to predict temperatures that are too low in fall, winter, and spring, but too high in summer. Dr. Beck further discusses the lack of coherence between Figures 14 through 21 and Figures 6 through 13 and its implications insofar as relying upon the model predictions.

The report is similarly deficient in explaining the meaning, significance, and limitations of the results in Figures 22-29. In essence, these figures report on the calibration of the stochastic model, plotting the results of equation 23 against those of equation 24. A goal of model calibration is to get these to match. Unfortunately, the key of these figures is insufficiently clear to distinguish which plotted line represents which result. The terminology of the figures deviates from that of the text, further confusing the results. As with the results in Figures 14-21, the results in Figures 22-29 show significant variations over time and, in at least some cases, a consistent mismatch.

Based on the comparisons in the figures, it is difficult to assess the quality of the calibration. More information on alternative calibration attempts would be helpful in this regard and would also give a sense of the model sensitivity. As well, segregation of the model-data comparison by month would help in identifying the accuracy with which the highest temperatures are predicted. Model predictions are particularly critical in this range because it is only this portion of the model results that are actually evaluated.

MODEL APPLICATION

On page 18, the report states goals that are not entirely congruent with the objectives stated on page 1. Also incongruent are the conclusions on page 20. While page 5 states that the model is a screening tool capable of identifying areas for further study, the report make no recommendations for further study. Instead, the report lists seemingly firm conclusions—an outcome that is inconsistent with the power and purpose of a screening model.

Model results are shown in terms of the frequency with which a temperature of 20°C is exceeded at the various stations along the rivers (in Figures 30-35) as well as the degree to which the temperatures are exceeded (in Figures 36-41). Simulation scenarios consider the current situation, the situation if existing dams were to be removed, and the situation if temperatures from tributary streams were kept less than or equal to 16°C. The simulations show that the frequency and magnitude with which 20°C is exceeded is decreased by removing dams (other than at the Snake River confluence and Grand Coulee Dam) and relatively unaltered by controlling tributary temperature.

There is a significant mismatch between the way the model was used and the way it was developed that calls into question its predictions. In its use, the model is applied only to evaluate extreme high temperatures that occur in the summer. But, the model's calibration and statistical evaluation were judged in terms of year-round agreement. The statistical measures used in the Kalman filtering evaluate the degree of agreement over the entire year and, for the deterministic model, via summations over the entire calibration period. Thus the summertime predictions, which tend to be high, are offset by the non-summer predictions, which tend to be low. The results presented in the report, however, show extreme temperature exceedances that occur only in the summertime period. Before the model can be confidently used to evaluate temperature extremes, it must be calibrated and checked specifically against periods of high temperature. Dr. Beck confirms this conclusion in his review.

This fundamental limitation notwithstanding, the model results predict temperature exceedances (in Figures 36 through 41) that are comparable to the calibration errors depicted in Figures 6 through 21. The "error bars" shown in Figures 36 through 41 may be confusing in this regard. They show the variation of the predicted exceedances around the mean and do not relate to the model uncertainty. However, it is clear from inspection of Figures 6 through 13 that the temperature model makes its poorest predictions at the extremes, yet it is precisely at the extremes where the model is being used.

SUMMARY

The EPA Columbia River temperature model uses unusual and technically sophisticated techniques to evaluate the effects of dams and other factors on temperature in the Columbia and Snake Rivers. Because an established model was not used, the Columbia River model should be verified in accordance with accepted practices for model quality assurance and quality control.

Calibration information provided for the model appears to show that the model predicts summertime temperatures that are generally higher than those observed and non-summer temperatures that tend to be lower than observed. However, the model calibration was evaluated in terms of year-round agreement, such that these two systematic errors balance each other. In contrast to the calibration evaluation, the model was used in a predictive mode only to evaluate extreme warm temperatures in the summertime. If the model is to be used primarily or solely to evaluate high temperature extremes, its predictive capability should be evaluated specifically for high temperature.

Errors in the model during summer appear to be comparable to the degree of exceedance predicted for summertime temperature excursion above the 20°C temperature threshold. This relative similarity of model error to the predicted excursion, as well as the mismatch between the calibration focus and prediction focus, indicate that the model results should be considered qualitative at best. As indicated in the report itself, the model is intended as a screening tool to identify areas for further research. As such, the model is not an adequate basis for policy decisions.

In a separate appended review, Dr. M. Bruce Beck focuses on the application of Kalman filtering in the EPA study. Dr. Beck concludes the Kalman filtering is implemented in a technically sound manner overall, but that certain aspects of the application require clarification. He also questions a number of explicit and implicit assumptions regarding the character of error and uncertainty and suggests additional analysis to explore their consequences.

REFERENCES

Brown, L.C., and T.O. Barnwell, 1987. The Enhanced Stream Water Quality Models QUAL2E and QUAL2E-UNCAS: Documentation and User Manual. Report No. EPA/600/3-87/007. U.S. Environmental Protection Agency, Athens, Georgia. May 1987.

Findikakis, A., F.A. Locher, and P.J. Ryan, 1980. Temperature and Turbidity Simulation in Spada Lake. In: H.G. Stefan, editor. *Proceedings of the Symposium on Surface Water Impoundments. June 2-5, 1980, Minneapolis, Minnesota.* Vol. I, Pg. 594-603. American Society of Civil Engineers, New York.

National Research Council, 1989. *Ground Water Models: Scientific and Regulatory Applications.* National Academy Press, Washington, D.C.

Shanahan, P., 1985. Water Temperature Modeling: A Practical Guide. in T.O. Barnwell, Jr., editor. Proceedings of the Stormwater and Water Quality Model Users Group Meeting, April 12-13, 1984. Report EPA-600/9-85-003. U.S. Environmental Protection Agency, Athens, Georgia.

TVA, 1972. Heat and Mass Transfer Between a Water Surface and the Atmosphere. Water Resources Research Laboratory No. 14. Tennessee Valley Authority, Division of Water Control Planning, Engineering Laboratory, Norris, Tennessee. April 1972.

Van Geer, F.C., C.B.M. Te Stroet, and Y. Zhou, 1991. Using Kalman Filtering to Improve and Quantify the Uncertainty of Numerical Groundwater Simulations, 1. The Role of System Noise and Its Calibration. *Water Resources Research*, Vol. 27, No. 8, Pgs. 1987-1994. August 1991.

Wunderlich, W.O., and R. Gras, 1967. Heat and Mass Transfer Between a Water Surface and the Atmosphere. Tennessee Valley Authority, Division of Water Control Planning, Engineering Laboratory, Norris, Tennessee.

APPENDIX

Review of Kalman filtering
by Dr. M. Bruce Beck

REVIEW OF
COLUMBIA RIVER TEMPERATURE ASSESSMENT: KALMAN FILTERING

INTRODUCTION

The purposes of this supplementary review are to:

(i) determine whether the Kalman filter has been used in a correct manner, i.e., in a manner consistent with published (and peer-reviewed) practice;

(ii) assess whether this filtering framework is appropriate for the given task;

(iii) consider alternatives to this framework suitable for any further such studies;

(iv) indicate what might be the desirable subjects of those future studies;

Before reporting on these matters, it will be helpful to review what has actually been achieved through applying the Kalman filter in this context.

GOAL OF EPA REPORT

Without seeking to diminish the significance of, or distract attention away from, variations in stream temperature over the entire annual cycle, let me suggest the real issue here is that of forecasting the maximum temperature over this cycle. Furthermore, if one is conservative in outlook, it may be better to over-predict than to under-predict this maximum. Re-stating this goal is important, because it has a bearing on some of the detail surrounding the way in which the filter has been used in the EPA Report.

"VALUE ADDED" THROUGH USE OF THE KALMAN FILTER

Besides the obviously highly charged political context of this problem, removing dams from the Columbia and Snake Rivers is a rather dramatic piece of policy. It would therefore seem important for all concerned to be reassured that such action is "right for the situation" and to be aware of the risks of "getting it wrong". Considerations of uncertainty and risk, then, are entirely appropriate in such a problem setting. Indeed, to have undertaken this exercise in the absence of such considerations, i.e., under the assumption of an entirely deterministic model, would itself have been an act of engaging in risk-taking (for the EPA). Use of the Kalman filter to address these issues of uncertainty is not usual, but by no means unknown.

The following is the essential role played by the filter in this study, to paraphrase (in perhaps colloquial terms):

> The world is uncertain. We know too that all models are approximations. All sources of uncertainty (approximations, omissions, errors) in the model will be subsumed under the label of the system noise vector ($\underline{w}$). Besides estimates of the model's conventional (deterministic) parameters, we shall therefore need estimates of the variance-covariance properties of $\underline{w}$ (the matrix Σ_Q) in order to account for the manner in which the inevitable residual uncertainty attaching to the model -- even when calibrated -- is propagated forward into forecasts of future behavior (under changed conditions).

In fact, looking at the source reference of van Geer *et al.* (1991), one might go so far as to say the primary purpose of calibration in the present study is to adjust the estimates of the variance-covariance matrix of the system noise (Σ_Q), with a view to assessing its impact on the uncertainty of the forecasts.

To be clear about what is subsumed under this matrix, we have the following generic sources of uncertainty:

(i) uncertainty in the (deterministic) parameters of the model;

(ii) uncertainty in the measured input disturbances of the model, i.e., here, principally the variations in the temperature of the tributaries;

(iii) uncertainty in all other unmeasured disturbances of the factors affecting temperature (the state variable).

In addition, account must be taken of uncertainty in the system's (past) observations, as must the uncertainty in the initial state of the system, i.e., the values of the spatial distribution of temperatures at the start of the calibration period and the forecasting period (although the author does not discuss this source of uncertainty). To be complete, we should also note that there will be a "structural error", or conceptual error, in the model. The manner in which the model's state variables interact with each other and the forms of the expressions used to describe these interactions will diverge from the (unknowable) "truth". There is currently no adequate method of accounting for errors of this form. This is hardly surprising: the problem is more philosophical than technical.

Given the decisions to account for uncertainty in this problem and to account for it using the Kalman filter, lumping the uncertainty in this manner under the single quantity (Σ_Q) is a pragmatic restriction, consistent with benefitting from the relative computational economy of the linear Kalman filter, when set against the alternative of Monte Carlo simulation, say. It also avoids having to use an extended Kalman filter, which would be necessary if one were to separate out (from Σ_Q) the parameteric uncertainty of the model, i.e., item (i) of the sources of uncertainty listed above (see also Beck and Halfon, 1991). The disadvantage of working with such an aggregate measure of uncertainty is that it will foreclose on any analysis of ranking the relative importance of the various sources of uncertainty, in terms of their contributions to the uncertainty of the predictions. Knowledge of this latter would be important in subsequently setting priorities for work that would be needed in order to reduce prediction uncertainty to some acceptable level (if it were thought to be unacceptably high for the purposes of making decisions).

PROPER USE OF THE KALMAN FILTER

As far as can be determined, Yearsley has used the Kalman filter -- for the purpose of calibration -- strictly in accordance with the procedures set out in the paper by van Geer *et al.* (1991). These authors, in their turn, make reference to the covariance-matching procedure of Mehra (1972), which, in spite of its vintage, remains the most common method for calibrating the variance-covariance properties of the system noise, i.e., for assigning values to the elements of the matrix Σ_Q. For the purpose of predicting the consequences of the policy options, again the filter has been applied in a manner consistent with normal practice (Beck and Halfon, 1991). To this extent, no fault can be found with the filter's application here; technically, the analysis appears to be sound.

There are, however, a number of places where caution should be exercised in interpreting the results of the Report. These are as follows.

(i) Figures 6 through 13 show comparisons of the simulated and observed water temperatures. Although we cannot be certain, it is quite probable that the corrected, or updated, estimate of the water temperature, i.e., $\underline{T}_k(+)$ from equation (9), has been used as the "simulated" value. If this is so, it is important to bear in mind that the results of these Figures may suggest a performance of the model better than what would have been achieved in the more familiar, purely deterministic setting, wherein the model is not embedded within a Kalman filter. Close inspection of equation (9) reveals the presence of the current observation of temperature $\underline{z}_k$. The effect of updating the one-step-ahead prediction, $\underline{T}_k(-)$, is thus always to draw any erroneous such prediction back towards the observation. The updated estimates $\underline{T}_k(+)$ reflect the benefit of this correction. To the eye trained on assessing a model's performance in the deterministic setting, without this "tracking" feature, a comparison of $\underline{T}_k(+)$ with the observation $\underline{z}_k$ can be deceptive. It might therefore be desirable to ask for clarification of whether the "simulated" values of Figures 6-13 represent $\underline{T}_k(+)$ or $\underline{T}_k(-)$.

(ii) Of the three policy options assessed (business-as-usual, removal of dams, control of tributary temperatures), the removal of the dams will clearly lead to a hydraulic regime unlike that of the (post-dam) observed record. The most obvious expectation of the consequences of this is that the uncertainty attaching to the hydraulic parameters estimated through the approximations of equations (13) through (15), if not any of the other (deterministic) model parameters, will be greater for this regime than for the presently observed conditions (with the dams in place). As far as can be established, no account is taken of this greater uncertainty; the same values of Σ_Q are used in generating the confidence bounds around all three sets of predictions. Since removal of the dams -- on the basis of the current analysis -- is predicted to have a significantly beneficial impact on lowering the number and magnitude of violations of the maximum temperature constraint, more detailed consideration of this point may well be warranted. Furthermore, the potential significance of this particular source of uncertainty may make it appropriate for future analyses to be based on explicit representation of the constituent sources of uncertainty, as opposed to their being lumped under Σ_Q.

(iii) It appears that the variance-covariance matrix of the system noise (Σ_Q) has non-zero elements on its leading diagonal alone, i.e., the assumption has been made that disturbances of the stream temperature dynamics are uncorrelated (primarily in space, it would appear). The Report is largely silent on the making of this assumption, although it is a common and not unreasonable one. Nevertheless, there is no discussion of its possible consequences, which is unfortunate since these may be material to the analysis. It is fairly widely appreciated that covariance among the elementary sources of uncertainty can have a significant effect on the propagation of uncertainty. In fact, it has generally been thought that it has the effect of reducing the degree of uncertainty attaching to the forecasts (this is not always the case, however; Beck and Halfon, 1991). We may note that van Geer *et al.* (1991) provide a means of assigning values to these off-diagonal elements of Σ_Q; it does not appear to have been used in the present analysis.

(iv) Comparing Figures 14 through 21 with respectively Figures 6 through 13 of the Report, is a surprisingly confusing task. If the principal issue at stake in this study is under-prediction of the maximum (summer) temperatures, it is especially important to be comfortable with the fact that the innovations ($\underline{\nu}_k$) are consistent with the relative positions of the quantities, $\underline{T}_k(+)$ (assumed) and $\underline{z}_k$, plotted in their respective Figures. Even after considerable reflection, I have failed to reconcile -- to my satisfaction -- the two sets of Figures.

To summarize, the subject of this review is a Report on a screening analysis designed to identify issues for further study. In general, the Kalman filter has been properly used for this purpose. However, the author of the Report has not identified all of the issues worthy of more detailed scrutiny.

APPROPRIATENESS OF FILTERING FRAMEWORK

In strategic terms, as already stated, it seems appropriate for uncertainty and risk to be parts of this assessment. In tactical terms, the Kalman filter provides (literally) a first-order approximation of error propagation. On balance this would appear commensurate with a preliminary screening analysis, although it is not common to find the Kalman filter employed in a study of this kind. In general, one could say the filter is often a good technique for problem discovery and definition, but one might subsequently want to apply some other form of analysis of the so defined subsequent problems.

Technically, if further use is to be made of the Kalman filter in assessing the Columbia river problem, it would be desirable to investigate the validity of assuming Gaussian distributions for the measurement errors and other sources of uncertainty. Significantly skewed distributions could compromise interpretation of the robustness of the predicted policy outcomes. Likewise, if (deterministic) parametric uncertainties are to be "unpacked" from the single aggregate (of the matrix Σ_Q), and a filtering framework remains the preferred computational setting, this could be achieved through the relatively minor extension of the extended Kalman filter (as in Beck and Halfon (1991)).

ALTERNATIVES

The obvious alternative to using the Kalman filter on a problem of this nature is Monte Carlo simulation, or some variation on that theme. Had this alternative been adopted, uncertainty would almost certainly have been accounted for in a different manner. In particular, as with virtually all Monte Carlo studies, the uncertainty attaching to the (deterministic) parameters of the model would have been the sole source of uncertainty accounted for. The question for calibration would then have been that of using the past observed temperatures in order to constrain, in some way, the choice of candidate parameterisations to be used for predicting the outcomes of the policy alternatives. Normally, one encounters Monte Carlo simulation in the context of forecasting (not model calibration). This requires specification of the statistical distributions to be used for the model's parameters, treated as random variables. In the absence of past observations, ranges of parameter values drawn from the literature are used to define these distributions. It is unusual to find studies using the set of past observations to generate "posterior" distributions of the parameters, for the purpose of forecasting, with the calibration process started with the "prior", literature-derived distributions.

In short, we derive models from uncertain theories reconciled with uncertain observations; we make predictions that are uncertain using models whose uncertainties will reflect all the successes and failures of calibration; and we must make decisions that are robust in the face of the resulting uncertain predictions, i.e., we must determine whether we would opt for the same course of action, all the uncertainties notwithstanding. Conceptually, the Kalman filter fits well with this view. If the alternative of Monte Carlo simulation were to be considered, it would probably find appropriate implementation through the procedure of Generalised Likelihood Uncertainty Estimation (GLUE) of Beven and Binley (1992).

POSSIBLE ISSUES FOR FURTHER STUDY

In the light of what has just been stated, regarding the account taken of uncertainty, from model development, through calibration and forecasting, into decision-making, the following could be of some significance. If one accepts the suggestion that the critical decision will turn on the reliability of the forecasts of maximum temperatures, then the manner in which the model is calibrated -- as the instrument of making this particular prediction -- should be geared to this goal. In practical terms, this implies that the covariance-matching technique employed for choosing Σ_Q should seek the best possible match over the periods of the summer maxima (as opposed to other seasons of the year, or over the year in some average manner). Figures 22 through 29 of the Report do not fully illuminate whether such a strategy has been pursued. We may probably conclude it has not.

Two criteria are used separately to rank the three policy alternatives, the number of days during the year when the temperature standard is exceeded and the magnitude of the excess temperature. It may be more meaningful to discriminate on the basis of a composite criterion, designed to capture the sense that the joint action of duration and magnitude of the excess is vital for the well-being of the endangered fish.

The option of removing the dams, in spite of the express consideration of uncertainty, still promises to bring about a significant change in the status quo. This is apparent from Figure 34 (when compared Figure 33) and, marginally more so, from the comparison of Figures 39 and 40. Making decisions under uncertainty -- as opposed to the determinism prevailing in its absence -- introduces greater subtlety (and complexity) into the debate. For example, in another context (Klepper *et al.*, 1991) the consequence of an action was forecast to have the effect of increasing the mean value of a commercial mussel culture, but also of introducing (relative to the status quo) a non-negligible risk of population collapse. While it is apparent that the present Report could have sustained such a more elaborate discussion, none is provided.

CONCLUSIONS AND RECOMMENDATIONS

This EPA Report, in my opinion, should contribute beneficially to the debate surrounding the survival of endangered species of fish in the Columbia River, precisely because of the way in which it casts its analysis in the setting of uncertainty and risk.

Although an unusual method to use, the Kalman filter has been implemented in a technically sound manner. On the whole the approximations and assumptions made in this implementation are consistent with the style of the investigation, this being that of a screening analysis. By implication, therefore, further study is likely to be needed before decisions on managing the thermal regime of the Columbia and Snake Rivers can be made.

Clarification should be sought on the following points: (i) the precise nature of the "simulated" values plotted in Figures 6 through 13; (ii) the possible impact on the predicted results of the policy alternatives of the likely higher uncertainties attaching to the model's hydraulic parameters in the event of removing the dams; (iii) the possible significance of covariance (as opposed to variance) among the sources of uncertainty accounted for in Σ_Q; and (iv) the consistency of interpretation of the results shown in Figures 14 through 21 relative to Figures 6 through 13.

If further study is to be undertaken by the EPA, one should seek to have the following issues addressed (among others raised in this review):

(i) a sensitivity analysis of the influence on prediction uncertainty of (a) an enlarged system noise variance-covariance matrix (Σ_Q), as a consequence of removing the dams, and (b) an altered set of values for the elements of this matrix as a result of gearing its calibration to the goal of matching covariances for the summer temperature maxima;

(ii) an assessment of prediction uncertainty when the specific sources of uncertainty are separated out from the aggregated form of Σ_Q, with a view to ranking the relative importance of these different sources;

(iii) an assessment of the normality of the distributions of various quantities manipulated through the filtering algorithms;

(iv) a more elaborate treatment of the implications of these, and any similar, subsequent, results for the debate surrounding decision-making under uncertainty.

REFERENCES

Beck, M B, and Halfon, E (1991), "Uncertainty, Identifiability and the Propagation of Prediction Errors: A Case Study of Lake Ontario", *J Forecasting*, **10**(1&2), pp 135-161.

Beven, K J, and Binley, A M (1992), "The Future of Distributed Models: Model Calibration and Predictive Uncertainty", *Hydrological Processes*, **6**, pp 279-298.

Klepper, O, Scholten, H, and van der Kamer, J P G (1991), "Prediction Uncertainty in an Ecological Model of the Oosterschelde Estuary", *J Forecasting*, **10**(1&2), pp 191-209.

Mehra, R K (1972), "Approaches to Adaptive Filtering", *IEEE Transactions on Automatic Control*, **10**, pp 693-698.

Van Geer, F C, Te Stroet, C B M, and Zhou, Y (1991), "Using Kalman Filtering to Improve and Quantify the Uncertainty of Numerical Groundwater Simulations. 1. The Role of System Noise and its Calibration", *Water Resources Research*, **27**(8), pp 1987-1994.

M B Beck

(b) (6) [redacted]

Athens

Georgia 30606-5734